Initiation of enzymatic replication at the origin of the *Escherichia* coli chromosome: Contributions of RNA polymerase and primase

(DNA/priming/oriC/plasmids)

TOHRU OGAWA*, TANIA A. BAKER, ARIE VAN DER ENDE[†], AND ARTHUR KORNBERG

Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305

Contributed by Arthur Kornberg, January 28, 1985

Replication of plasmids that depend on the 245-base-pair origin of the Escherichia coli chromosome (oriC) requires many purified proteins that (i) direct initiation to oriC (e.g., dnaA protein), (ii) influence initiations elsewhere (e.g., auxiliary proteins), and (iii) prime and extend DNA chains (e.g., priming and synthesis proteins). For the RNA priming and initiation of new DNA chains, the requirements for both primase and RNA polymerase (RNA pol) [Kaguni, J. M. & Kornberg, A. (1984) Cell 38, 183-190] have been further analyzed. Depending on the levels of auxiliary proteins (topoisomerase I and protein HU), three priming systems can operate: primase alone, RNA pol alone, or both combined. At low levels of auxiliary proteins, primase alone sustains an effective priming system. At higher levels, primase action is blocked, but RNA pol alone can initiate replication, albeit feebly; at these high levels of auxiliary proteins, primase and RNA pol act synergistically. When RNA pol is stalled by an inhibitor or lack of a ribonucleoside triphosphate, primase action is also inhibited. Based on these and other data [van der Ende, A., Baker, T. A., Ogawa, T. & Kornberg, A. (1985) Proc. Natl. Acad. Sci. USA 82, in press], RNA pol can counteract inhibition by auxiliary proteins and thus activate the origin for the priming by primase of the leading strand of the replication fork.

Replication of plasmids in Escherichia coli that depend on the chromosomal origin of replication (oriC) has been partially reconstituted with purified proteins (1-6). These proteins fall into three functional groups: initiation proteins (e.g., dnaA protein) for recognition of the oriC sequence and assembly of a replication complex, specificity (auxiliary) proteins for suppression of initiation of replication at sites other than oriC, and replication proteins for priming and elongation of DNA chains. A key event in the initiation process is specific binding of dnaA protein to the oriC sequence (2, 7). RNA polymerase (RNA pol) involvement in initiation of replication of the E. coli chromosome in vivo is based on inhibition of a new cycle of replication by rifampicin (8), suppression of dnaA mutations by rpoB mutations (9, 10), and increased chromosomal copy number in rpoB or rpoC mutants (11, 12). In oriC replication in vitro, the RNA pol-dependent event precedes DNA synthesis (1, 6), but its precise role has not been clarified. In the absence of primase, RNA pol was presumed to prime the synthesis of the leading strand at oriC

Under certain conditions examined in this and a companion study (13), DNA synthesis takes place in the absence of RNA pol, whereas primase is essential, indicating a role of RNA pol in activating the origin for initiation of replication as proposed for replication of bacteriophage λ DNA (14). The present study examines various parameters, particularly the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

auxiliary proteins, to evaluate the contributions of RNA pol and primase to initiation of oriC replication. Three priming systems have been revealed: primase alone, RNA pol alone, and both combined. Further characterization of the solo primase system is presented elsewhere (13). These studies establish the capacity of primase to prime the leading strand of DNA replication and point to a role for RNA pol in activating the oriC sequence as the unique initiation site.

MATERIALS AND METHODS

Strains, Phages, and DNAs. E. coli CM987 (F⁻, thi, asn, recA) and oriC plasmid pCM959 (15) were generously provided by M. Meijer (University of Amsterdam, The Netherlands). pCM959 DNA (see below) was prepared by alkaline lysis of CM987 cells that harbor the plasmids (16). Phages M13oriC26 (17) and M13\Delta E101 (18) were from J. M. Kaguni and D. S. Ray, respectively. Covalently closed supercoiled replicative form (RFI) DNA of these phages was prepared from infected K37 (Hfr) cells as described (19). The DNAs were further purified by phenol/chloroform treatment, two successive bandings in CsCl/ethidium bromide density gradients, ethanol precipitation, and gel filtration through Bio-Gel A-5m.

Reagents and Enzymes. Sources were ribonucleoside triphosphates (rNTPs) and phosphocreatine from Sigma; deoxyribonucleoside triphosphates (dNTPs) from P-L Biochemicals; $[\alpha^{-32}P]dTTP$ (>800 Ci/mmol; 1 Ci = 37 GBq) from Amersham; rifampicin from Calbiochem; streptolydigin, a gift from Upjohn; highly purified DNA replication proteins as described (6).

Reconstitution Assay for DNA Replication. The novel requirement of oriC replication in a crude enzyme fraction for a high concentration of a hydrophilic polymer, such as polyvinyl alcohol or polyethylene glycol (20), has been replaced in a reconstituted system by an increased level of the gyrase A subunit (6) and the removal of inhibitors (unpublished). In such a reconstituted system, stimulation of replication by crude enzyme fractions, such as flowthrough and eluate fractions from an Amicon red-A agarose column (1, 6), is no longer significant (data not shown). Thus, maximal DNA synthesis is achieved by a combination of highly purified enzymes (see below). The minichromosome pCM959 (15), which consists solely of E. coli DNA encompassing oriC (bp -677 to bp +3335, where bp = base pair) (21), was used as the oriC DNA template, except where indicated, to avoid possible complications by sequences of the vector portion of

Abbreviations: RFI, covalently closed supercoiled replicative form; rNTP and dNTP, ribo- and deoxyribonucleoside triphosphate; SSB, single-stranded DNA binding protein; bp, base pair; RNA pol, RNA polymerase; Topo I, topoisomerase I; Pol III, DNA polymerase III. *Present address: Institute of Molecular Biology, School of Science,

Nagoya University, Chikusa, Nagoya 464, Japan.

†Present address: Department of Cell Biology, Medical Faculty, University of Utrecht, Utrecht, The Netherlands.

previously used M13-oriC chimeric phage DNA (1-6) (unpublished).

The standard reaction (25 µl) contained Hepes·KOH (pH 7.6), 40 mM; magnesium acetate, 8 mM; phosphocreatine, 2 mM; dithiothreitol, 4 mM; ATP, 2 mM; GTP, CTP, and UTP, 500 μ M each; dATP, dGTP, and dCTP, 100 μ M each; $[\alpha^{-32}P]dTTP$, 100 μ M, 200–600 cpm/pmol; supercoiled template DNA, 200 ng (600 pmol of nucleotide); bovine serum albumin, 0.4 mg/ml; creatine kinase, 300 ng; protein HU, 20 ng; topoisomerase I (Topo I), 10 ng; RNase H, 1.4 ng; gyrase A subunit, 360 ng; gyrase B subunit, 500 ng; RNA pol, 400 ng; dnaA protein, 110 ng; single-stranded DNA binding protein (SSB), 320 ng; dnaB protein, 100 ng; dnaC protein, 50 ng; primase, 16 ng; and DNA polymerase III (Pol III) holoenzyme, 100 ng. Primase was omitted in the solo RNA pol reactions and RNA pol was omitted in the solo primase reactions. The mixtures were assembled at 0°C and incubated for 30 min at 30°C. Total nucleotide incorporation was measured in a liquid scintillation counter after trichloroacetic acid precipitation onto Whatman GF/C glass-fiber filters.

RESULTS

Replication Can Occur in the Absence of Either RNA pol or Primase. Upon varying the levels of the reaction components, significant replication was consistently observed when either RNA pol or primase was omitted (Table 1). The solo primase system appeared far more efficient than the solo RNA pol system; yet they have the same requirements for the other replication proteins.

Topo I and a High Level of Protein HU Inhibit the Solo Primase System. As auxiliary factors, the effects of Topo I and protein HU on the solo primase system were most striking (Figs. 1 and 2). Although high levels of Topo I were generally inhibitory, its action at lower levels [effective for specificity function at oriC (5)] was most profound on the solo

Table 1. Requirements for replication in reactions with RNA pol. primase, or both

!	DNA synthesis, pmol			
Component omitted	RNA pol	Primase	RNA pol + primase	
None	54	489	853	
Protein HU	10	184	287	
dnaA protein	3	3	5	
Gyrase A subunit	14	6	49	
Gyrase B subunit	1	1	3	
SSB	7	28	41	
dnaB protein	1	1	3	
dnaC protein	1	2	3	
Pol III holoenzyme	0	0	1	

Complete reaction contains Topo I and RNase H to maintain specificity for oriC DNA and dependence on dnaA protein. Incomplete dependence of incorporation on protein HU and gyrase A subunit is partly due to the contamination of small amounts of these proteins in the other protein fractions. Less than 5 pmol of incorporation was observed when both RNA pol and primase were omitted from the complete reaction mixture.

primase system. This inhibitory effect was due in large part to prolonging the lag time of the reaction (Fig. 2) (13).

Protein HU, which stimulated all reactions when present at a low level, became inhibitory at elevated levels, especially on the solo primase system (Fig. 1). At twice the fully inhibitory level, protein HU (at a 1:1 weight ratio to DNA) converts form I of simian virus 40 into a nucleosome-like, beaded structure in the presence of a type I topoisomerase (22) and enhances transcription of phage λ DNA (23). At an optimal level of protein HU (1:20 weight ratio to DNA), the solo primase system appeared to be as efficient as the one containing both primase and RNA pol. In contrast, the system lacking primase was inefficient at all concentrations

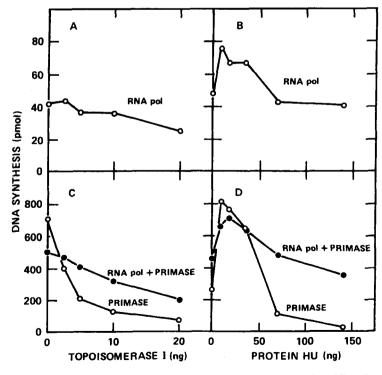


FIG. 1. Effect of Topo I and protein HU on the reactions with RNA pol, primase, or both. Titration of Topo I was performed in the presence of 20 ng of protein HU. Titration of protein HU was performed in the absence of Topo I.

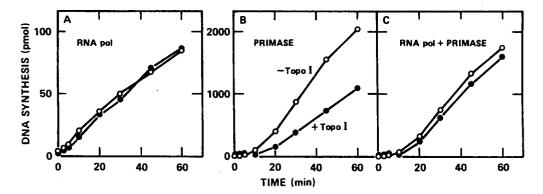


Fig. 2. Time course of replication and the influence of Topo I. Reaction mixtures (225 µl) of the three priming systems were assembled in the presence (•) or absence (o) of Topo I. DNA synthesis was measured on 25-μl samples taken at the indicated times of incubation at 30°C.

of Topo I and protein HU. The strong dependence on RNA pol observed previously in reconstituted systems (6) was due in large part to the high levels of Topo I and protein HU that were used.

The Initiation Systems Depend on dnaA Protein and oriC DNA. RNase H (4) and Topo I (5) have been identified as specificity proteins that suppress DNA synthesis that is not dependent on dnaA protein and the oriC sequence. The solo primase system did not require these auxiliary proteins to maintain dnaA protein dependence of replication (Table 2) nor was replication observed when DNA templates lacked the oriC sequence (Table 3). On the other hand, a low level of DNA synthesis in the absence of dnaA protein was detected in reactions that contained RNA pol but lacked auxiliary proteins, particularly RNase H (Table 2). DNA synthesis was also observed with DNA lacking the oriC sequence (to an extent depending on the template) in reactions that contained RNA pol but did not contain auxiliary proteins (Table 3). The contribution to the total DNA synthesis of this nonspecific DNA synthesis was most significant in the solo RNA pol system, where only a small amount of dnaA protein- and oriC-dependent replication was observed. RNase H had no effect on the dnaA proteindependent DNA synthesis in the solo primase system but stimulated the dnaA protein-dependent reaction by 50-100% when both RNA pol and primase were present (data not shown). The stimulation by auxiliary proteins of the combined RNA pol and primase reaction with the oriC DNA template (Table 3) was due to this activity of RNase H.

Inhibition of the solo primase reaction with M13oriC26 by auxiliary proteins was especially great (Table 3), presumably due to a prolonged lag time caused by Topo I on this very large template (13).

Solo RNA pol System Is Inefficient. In contrast to the solo primase system, the solo RNA pol system was rather feeble and could not be enhanced by systematic adjustments of the reaction components. To insure that this limited level of

Table 2. Effect of auxiliary proteins on replication of pCM959 DNA in the absence of dnaA protein

Auxiliary proteins	dnaA protein	DNA synthesis, pmol			
		RNA pol	Primase	RNA pol + primase	
None	+	54	489	853	
None	_	47	. 5	43	
Торо I	_	31	4	27	
RNase H	_	3	3	5	
Topo I and RNase H	-	3	3	5	

replication was not due to traces of primase in the other reagents, antibody against primase was shown to selectively inhibit primase-dependent replication but not to affect the RNA pol-dependent reaction (Table 4). Thus, RNA pol can prime DNA synthesis under certain conditions that depend on the oriC sequence, dnaA protein, and other key replication proteins. However, the inefficiency of this reaction suggests that primase is an essential component of the replication machinery at a growing fork and not replaceable by RNA pol.

rNTPs, Other Than ATP, Are Not Required in the Solo Primase System. The amount of replication observed upon omission of one or more of the rNTPs was strikingly different in the three initiation systems (Table 5). The solo RNA pol reaction was totally dependent on all four rNTPs, while the solo primase reaction required only ATP and was inhibited by adding rGTP, rCTP, and rUTP; ATP is essential for the functions of gyrase, dnaB protein and Pol III holoenzyme, as well as primase. This result is consistent with the capacity of primase to prime DNA synthesis on the SSB-coated singlestranded circular DNA of bacteriophage M13 in the presence of only ATP and the dNTPs (24). The enhanced replication, in the absence of rGTP, rCTP, and rUTP, suggests that extensive primer synthesis competes with use of the template for efficient replication.

In contrast to the solo primase system, the reaction that contained both RNA pol and primase was inhibited by

Table 3. Comparison of template DNAs

Template (size, kb)		DNA synthesis, pmol			
	Auxiliary proteins	RNA pol	Primase	RNA pol + primase	
pCM959	_	163	1095	553	
(4.0)	+	41	409	811	
M13oriC26	_	173	665	279	
(12.3)	+	68	63	351	
φX174	_	115	6	174	
(5.4)	+	7	4	8	
pBR322	-	65	10	84	
(4.4)	+	10	7	15	
M13ΔE101	_	168	29	269	
(6.3)	+	8	16	20	

Reactions were performed with or without auxiliary proteins (Topo I and RNase H). Only pCM959 and M13oriC26 contain the minimal oriC sequence, flanked by only E. coli DNA in pCM959 and by E. coli, G4, and M13 sequences in M13oriC26; φX174 is the phage replicative form (RFI); pBR322 is a ColE1-like plasmid; M13AE101 (lacking the origin for complementary strand synthesis) and M13oriC26 are RFI DNAs. kb, Kilobase pairs.

Table 4. Effect of primase antibody

Reaction	RNA pol,	Primase, ng	Antiserum, µl	DNA synthesis pmol
RNA pol	400	0	0	35
	400	0	0.025	40
	400	0	0.05	42
Primase	0	16	0	502
	0	16	0.025	23
	0	16	0.05	12
	0	40	0.025	441
RNA pol	400	16	0	470
+ primase	400	16	0.025	45
	400	16	0.05	38
	400	40	0.025	460

Reactions were performed with the indicated amounts of RNA pol, primase, and antiserum (prepared in mice with DEAE-purified primase and heated 15 min at 56°C).

omitting rCTP, rGTP, and rUTP, especially at higher levels of RNA pol (Fig. 3). Similarly, replication was blocked when RNA pol was inhibited by rifampicin or streptolydigin (Fig. 3). Thus, the primase reaction is blocked when RNA pol is inhibited. When core RNA pol (lacking the σ factor) was used in place of RNA pol holoenzyme, no significant inhibition was observed either upon omission of rGTP, rCTP, and rUTP or by addition of rifampicin or streptolydigin (data not shown). These observations may be explained by RNA pol being fixed at a promoter(s) in the presence of an inhibitor or stalled for lack of rNTPs and thus blocking a stage in replication. Inhibition of transcription by an RNA pol-rifampicin complex bound to a promoter has been reported (25). When RNA synthesis by RNA pol was allowed to proceed (for 10-20 min), then stopped by rifampicin and streptolydigin, while at the same time DNA synthesis was started by addition of dNTPs, replication was not inhibited (ref. 6; data not shown). These findings suggest that the role of RNA pol is limited to the initiation stage of the reaction.

DISCUSSION

More than 10 purified proteins are needed to reconstitute the specific replication of plasmids that depend on the unique origin (oriC) of the E. coli chromosome (6). Among these proteins are some that direct initiation of DNA strands to the oriC sequence and others (auxiliary, specificity proteins) that suppress initiation elsewhere. In this study, evaluating the contributions of RNA pol and primase to the priming of new DNA chains, three systems have been revealed: (i) solo primase, (ii) solo RNA pol, and (iii) primase and RNA pol combined.

The solo primase system can support efficient replication provided that the auxiliary protein, Topo I, is absent or at a low level and that protein HU is at a weight ratio to DNA of about 1:20; at a ratio of 1:2, protein HU inhibits primase

Table 5. Requirement for rNTPs

rNTPs	DNA synthesis, pmol			
	RNA pol	Primase	RNA pol + primase	
A, G, C, U	84	241	673	
- G, C, U	3	3	5	
A, - C, U	7	685	173	
A, G,	9	448	128	
A,	6	1045	191	

Reactions were performed as described in *Materials and Methods* except for omission of rNTPs as indicated.

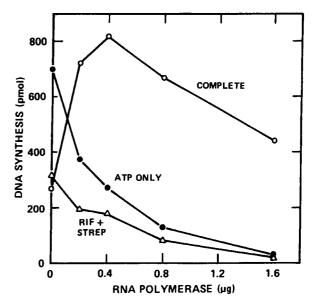


Fig. 3. Effect of omission of rGTP, rCTP, and rUTP or addition of rifampicin and streptolydigin on the reaction containing primase and various amounts of RNA pol. Complete reactions were performed as described in *Materials and Methods* except that the amounts of RNA pol were changed. In the "ATP ONLY" reactions, rGTP, rCTP, and rUTP were omitted. In the "RIF + STREP" reactions, 20 μ g of rifampicin per ml and 200 μ g of streptolydigin per ml were added. Similar results were obtained with either rifampicin alone or streptolydigin alone.

action completely (Fig. 1). Neither these nor other auxiliary proteins (e.g., RNase H) are needed to insure oriC specificity and the absolute dependence of replication on dnaA protein bound to the oriC sequence (7) (Tables 2 and 3). Further characterization of the solo primase system and divisibility of initiation into prepriming and priming stages are presented elsewhere (13).

In contrast to the solo primase system, the solo RNA pol system is only 1/10th as active and is relatively unaffected by the auxiliary proteins, except when they are needed to suppress initiation at non-oriC regions of an oriC template or on supercoiled DNAs that do not contain oriC (Fig. 1, Table 3)

In the system that includes both RNA pol and primase, the RNA pol relieves much of the inhibition of the primase system imposed by high levels of Topo I and protein HU but has little effect on replication at low levels of these auxiliary proteins. However, the auxiliary proteins are needed for specificity to block the use of non-oriC DNA whenever RNA pol is present. When RNA pol transcription is inhibited by rifampicin or streptolydigin or stalled for lack of a rNTP, initiation of replication by either primase or RNA pol is prevented (Fig. 3). These inhibitory effects are not observed once transcription has taken place.

Reactions containing primase show a lag of 10-15 min, whereas the lag time of the solo RNA pol reaction is only 2-5 min (Fig. 2). All reactions proceed for >60 min. Prior incubation of the primase-dependent reaction mixtures in the absence of dNTPs removes the lag and a burst of replication is observed upon addition of dNTPs (6, 13). Prior incubation of the solo RNA pol reaction mixture, on the other hand, results in only a small reduction in the lag time (data not shown). When an assay system is used that allows only a limited period for initiation and a single round of elongation, only 10-20% of the input DNA molecules serve as templates in reactions that contain primase, while even fewer (<5%) are used in the solo RNA pol reaction (ref. 13; data not shown).

Thus, several features distinguish the primase-independent reaction from those that contain primase.

These and other findings (13) suggest that primase has a major role in priming the leading and lagging strands of replication forks. Although RNA pol transcription can also prime DNA synthesis, it is far less efficient than primase and may make its chief contribution by activating the *oriC* region for priming by primase. Possibly, a factor involved in determining the frequency of initiation of replication of the E. coli chromosome could control the synthesis of RNA that activates the origin.

How the properties of primase and RNA pol in these reconstituted reactions apply to physiological conditions is difficult to judge. RNA pol has been reported to participate in initiation of replication of the *E. coli* chromosome (8–12, 20). However, inferences of involvement of RNA pol based on actions of inhibitors (e.g., rifampicin) in cells (8) and crude enzyme systems (20) may be unsound because of the capacity of an inhibitor–RNA pol complex to block initiations by primase as well (Fig. 3). Independence from RNA pol in replication of R1 plasmids has been demonstrated (26) and implies a reliance on primase for priming this double-stranded template.

Despite the multiplicity of proteins in the replication system, it seems likely that the true oriC system demands still more. For example, several partially purified fractions are known to stimulate the rate or extent of replication (1, 6). These fractions were not replaced by available replication proteins, such as proteins n, n', n", and i, dnaJ protein, dnaK protein, rep protein, helicase II, DNA ligase, and DNA polymerase I.

We thank LeRoy L. Bertsch for expert assistance in preparing this manuscript. This work was supported by grants from the National Institutes of Health and the National Science Foundation. A.v.d.E. was supported by the Netherlands Organization for the Advancement of Pure Research (Zuiver Wetenschappelijk Onderzoek) with a North Atlantic Treaty Organization Science Fellowship.

- Fuller, R. S., Bertsch, L. L., Dixon, N. E., Flynn, J. E., Jr., Kaguni, J. M., Low, R. L., Ogawa, T. & Kornberg, A. (1983) Mechanisms of DNA Replication and Recombination: UCLA Symposia on Molecular and Cellular Biology, New Series, ed. Cozzarelli, N. (Liss, New York), Vol. 10, pp. 275-288.
- Fuller, R. S. & Kornberg, A. (1983) Proc. Natl. Acad. Sci. USA 80, 5817-5821.

- Dixon, N. E. & Kornberg, A. (1984) Proc. Natl. Acad. Sci. USA 81, 424-428.
- Ogawa, T., Pickett, G. G., Kogoma, T. & Kornberg, A. (1984) Proc. Natl. Acad. Sci. USA 81, 1040-1044.
- Kaguni, J. M. & Kornberg, A. (1984) J. Biol. Chem. 259, 8578-8583.
- 6. Kaguni, J. M. & Kornberg, A. (1984) Cell 38, 183-190.
- Fuller, R. S., Funnell, B. E. & Kornberg, A. (1984) Cell 38, 889-900.
- 8. Lark, K. G. (1972) J. Mol. Biol. 64, 47-60.
- Bagdasarian, M. M., Izakowska, M. & Bagdasarian, M. (1977)
 J. Bacteriol. 130, 577-582.
- Atlung, T. (1981) ICN-UCLA Symp. Mol. Cell. Biol. 22, 297-314
- Rasmussen, K. V., Atlung, T., Kerszman, G., Hansen, G. E. & Hansen, F. G. (1983) J. Bacteriol. 154, 443-451.
- Tanaka, M., Omori, H. & Hiraga, S. (1983) Mol. Gen. Genet. 192, 51-60.
- van der Ende, A., Baker, T. A., Ogawa, T. & Kornberg, A. (1985) Proc. Natl. Acad. Sci. USA 82, in press.
- Dove, W. F., Inokuchi, H. & Stevens, W. F. (1971) The Bacteriophage Lambda, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 747-771.
- von Meyenberg, K., Hansen, F. G., Riise, E., Bergmans, H. E. N., Meijer, M. & Messer, W. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 121-128.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 90-91.
- Kaguni, J. M., LaVerne, L. S. & Ray, D. S. (1979) Proc. Natl. Acad. Sci. USA 76, 6250-6254.
- Kim, M. H., Hines, J. C. & Ray, D. S. (1981) Proc. Natl. Acad. Sci. USA 78, 6784-6788.
- Davis, R. W., Botstein, D. & Roth, J. R. (1980) Advanced Bacterial Genetics (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 116-117.
- Fuller, R. S., Kaguni, J. M. & Kornberg, A. (1981) Proc. Natl. Acad. Sci. USA 78, 7370-7374.
- 21. Buhk, H.-J. & Messer, W. (1983) Gene 24, 265-279.
- Rouvière-Yaniv, J., Yaniv, M. & Germond, J.-E. (1979) Cell 17, 265-274.
- Rouvière-Yaniv, J. & Gros, F. (1975) Proc. Natl. Acad. Sci. USA 72, 3428-3432.
- 24. Rowen, L. & Kornberg, A. (1978) J. Biol. Chem. 253, 770-774.
- Kassavetis, G. A., Kaya, K. M. & Chamberlin, M. J. (1978) Biochemistry 17, 5798-5804.
- 26. Diaz, R. & Ortega, S. (1984) Nucleic Acids Res. 12, 5175-5191.